Randomised trialomania? The multicentre liver transplant trials of tacrolimus

Thomas E Starzl, Allan Donner, Michael Eliasziw, Larry Stitt, Paul Meier, John J Fung, John P McMichael, Satoru Todo

Randomised trialomania? The multicentre liver transplant trials of tacrolimus

Thomas E Starzl, Allan Donner, Michael Eliasziw, Larry Stitt, Paul Meier, John J Fung, John P McMichael, Satoru Todo

Introduction

In 1989–90 our group reported that tacrolimus (FK 506) could systematically control liver allograft rejection that had been intractable despite conventional cyclosporin-based immunosuppression.[1,2] The "rescued" patients were maintained on tacrolimus, and manifested no unique or unexpected toxicity. Consequently, a pilot trial was begun in which tacrolimus was substituted for cyclosporin from the time of operation. By early 1990, nearly 200 liver, kidney, and other kinds of organ recipients who had been entered had superior actuarial survival, lower requirement for prednisone, and better quality of life than observed in our past experience.[3,4] The upgrading of outlook after liver transplantation[3] was as obvious as when cyclosporin succeeded azathioprine as the baseline immuno-suppressant a decade before.[5] By November, 1993, 1391 primary liver-transplant recipients had been treated in Pittsburgh with the new drug.[6] Only 35 (2·5%) of the patients crossed over from tacrolimus to cyclosporin and 15 of these changed back when rejection supervened.

The keystone management principle

The new drug was user friendly. As with all previous baseline immunosuppressants,[7–9] the management secret with tacrolimus was administration of doses up to the limit imposed by the drug's toxicity. Dose-manoeuvrable prednisone or other adjuvant agents were then added as needed to control or reverse rejection, or given prophylactically. Because the dose-limiting side-effects of cyclosporin[10] and tacrolimus[3,4,11,12] were the same (nephrotoxicity, neurotoxicity, and diabetogenicity), these organ-system-specific toxic manifestations could be used from the first day of treatment to determine dose ceilings; the occurrence of rejection helped to establish the floor. The folly of making invidious toxicity comparisons between cyclosporin and tacrolimus when the scales could be tilted one way or the other by ratcheting the doses up or down was self-evident.[12] The only adverse effects

observed exclusively with one but not the other drug were dose-related hirsutism, gingival hyperplasia, and facial brutalisation with cyclosporin.[10]

It was easy, as it had been a decade earlier with cyclosporin,[13] to deduce the meaning of trough plasma and blood concentrations (the plasma/blood ratio was about 0·1) and to relate these to toxic manifestations, rejection, and the preceding tacrolimus dose. When we realised that the first Pittsburgh patients had been overdosed, this was corrected in subsequent cases within a few postoperative days or hours by responding to the clinical events with flexible dosing. Nevertheless, we had lowered both the starting intravenous and oral doses in Pittsburgh by January, 1990;[12] subsequently these were reduced again.[14]

These were important revisions no matter what the transplanted organ but especially so with the liver because the metabolism of tacrolimus is more dependent than that of cyclosporin on good hepatic function.[12,15,16] In addition, absorption of tacrolimus is little disturbed compared with that of cyclosporin by the absence of bile or by intestinal disorders.[15] These and other details of the pharmacokinetics of tacrolimus, dose ranges, appropriate management strategies, and adverse events were well worked through by the time of meetings to organise multicentre trials in March, 1990. 3 weeks later, the same data were presented to the American Surgical Association (on April 5, 1990) and a manuscript was published 5 months later on the eve of the multicentre trials.

The Pittsburgh trials and beyond

By the time the multicentre investigations began in late August (North America) and September (Europe), 1990, our Pittsburgh randomised trial was more than half-finished.[14] This single-centre trial followed a crisis that erupted during the summer and autumn of 1989 in our centre and at the US Food and Drug Administration (FDA). The problem was caused by the mass referral to Pittsburgh from other centres of dozens of patients needing tacrolimus rescue therapy. Understandably, the drug was being demanded by directors of other programmes. The only investigational new drug (IND) licence in existence at the time had been issued to TES, who could not legally distribute tacrolimus secondarily. Moreover, Fujisawa (the manufacturer) had not yet decided whether to proceed with the drug's development because of reports from Europe of toxicity in animals. At our request and with the support of Fujisawa, which expedited its own IND, the FDA placed tacrolimus on the fast-track for evaluation. However, the extent of testing required to establish safety and efficacy became

Pittsburgh Transplantation Institute, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA (Prof T E Starzl MD, J J Fung MD, J P McMichael BSc, Prof S Todo MD); **Department of Epidemiology and Biostatistics, University of Western Ontario, Ontario, Canada** (Prof A Donner PhD, M Eliasziw PhD, L Stitt MSc); **and Department of Statistics, Columbia University, New York City, New York** (Prof P Meier PhD)

Correspondence to: Prof Thomas E Starzl, 3601 Fifth Avenue, 4C Falk Clinic, Pittsburgh, PA 15213, USA

contentious. A controlled-rescue trial that would abandon half the patients to control status was not a savoury prospect. In addition, pressure had begun to mount to force tacrolimus to jump through hoops similar to (although more constrictive than) those negotiated by cyclosporin nearly 10 years earlier.

Eventually, an FDA advisory committee recommended two multicentre randomised trials that had little bearing on the life-saving rescue capability of tacrolimus that qualified it for the FDA fast-track. Tacrolimus and cyclosporin, given from the time of liver transplantation, would be compared head to head. At FDA request, the protocol shown in figure 1 was prepared in Pittsburgh for this purpose, formally submitted to the FDA in November, 1989, and approved 1 month later. The document was a product of administrative pressure, and was written in a climate of anxiety. The ethical quandary of randomising treatment to "life stake" liver transplant recipients was similar to that encountered when cyclosporin replaced azathioprine as the principal immunosuppressant.[17] Realising that we considered the superiority of tacrolimus to be settled, the FDA granted the Pittsburgh team's request to be excused from participation.[18]

The randomisation protocol was implemented at the Presbyterian University Hospital (Pittsburgh) in February, 1990,[14] at the insistence of the institutional review board that governed clinical research at this facility. However, the two independent institutional review boards at the university's paediatric and veteran's hospitals permitted tacrolimus to be used by patient's/physician's choice. The bitter split within the university came from the recognition that a randomised trial was not ethical in the absence of equipoise—ie, in the absence of substantial uncertainty about the relative benefit of the treatments compared. At the insistence of the clinicians caring for the patients, a multi-institutional "patients' rights committee" (consisting of faculty members at University of Pittsburgh, Carnegie Mellon University, and Harvard Medical College) was convened to evaluate the results every 3 months and make recommendations to the institutional review board about continuation or stoppage of the trial.

The safety and efficacy comparison of the competing drugs was ensured by equalisation at the outset of all treatment variables except the competing drugs (figure 1) and by clear definition of end-points. More importantly, the patients were protected by allowing them early access to whichever drug had the better therapeutic index. By the time the trial was prematurely closed in 1991 at the recommendation of the oversight committee, 154 patients had been randomised. The result, analysed by intention to treat, will be reported in detail elsewhere. Massive crossover from cyclosporin to tacrolimus left only 24 patients on cyclosporin after the first year. There was only
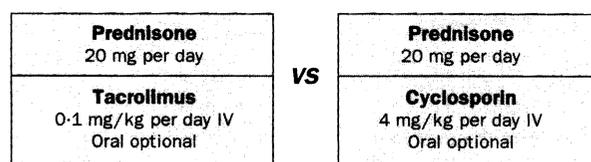

Figure 1: **Balanced experimental design in Pittsburgh randomised trial**
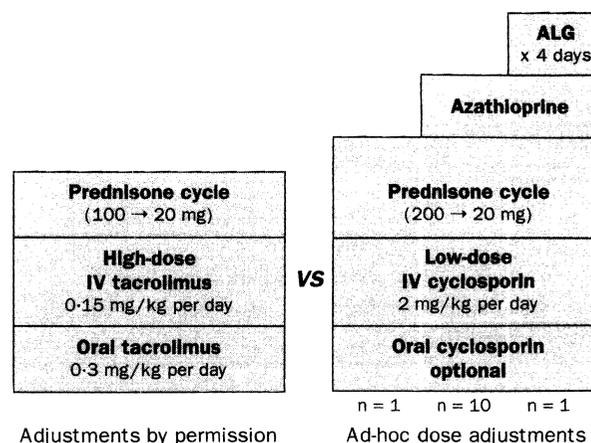IV=intravenous.

Figure 2: **Unbalanced experimental design in multicentre trials[19,20]**
American details[20] shown. ALG=antilymphocyte globulin. n=number of centres.

one crossover from tacrolimus to cyclosporin. Throughout the entire $3\frac{1}{2}$–5-year period of study, tacrolimus enjoyed a statistically significant (log rank) greater freedom from rejection alone or in combination with freedom from graft loss and adverse events.

The multicentre randomised trials

The FDA-approved Pittsburgh protocol (figure 1) was presented and recommended by TES at both multicentre organisational meetings in early March, 1990. It was rejected in favour of the design depicted in figure 2. For the American trial, the cyclosporin arm was uploaded with twice the induction doses of prednisone in all twelve centres, a third drug (azathioprine) in eleven, and a fourth agent (polyclonal ALG) in one. The eight European centres also had similar unbalanced and diverse protocols.

On both sides of the Atlantic, the starting and all subsequent cyclosporin doses were left to physician's discretion. By contrast, the stipulated starting doses of tacrolimus were 50% higher than those being used in Pittsburgh. Post-transplant adjustments were to be largely contingent on blood-monitoring results that would not be available until several days later because the samples were to be shipped for analysis to reference laboratories in distant cities. Formal starting-dose revisions were not made until 30% and 18% of the European and American tacrolimus cases had been enrolled. The combination of excessive dosage and sluggish response time to toxic events in the tacrolimus arm had devastating consequences. The trials were saved by the skill of the investigators who overrode the protocol's dose restrictions.

The results (as in our Pittsburgh randomised trial) were analysed by intention to treat, which protects the original randomisation by crediting the end-point outcome to the assigned treatment, even in the presence of protocol violations. The only legitimate reasons to discontinue compilation of end-point data (called censoring or termination) were patients' death or loss to follow-up.

The European report

A 5% better survival of patients was recorded in the tacrolimus arm (46 vs 61 deaths) and a 5% higher graft survival.[19] The survival advantage was not statistically significant, but about 10% of the surviving grafts credited

	Tacrolimus	Cyclosporin
Randomised	263	266
Total censored	83 (31·6%)	102 (38·3%)
Reasons for censoring		
Death	14 (5·3%)	16 (6·0%)
2nd transplantation for technical problem	17 (6·5%)	21 (7·9%)
Adverse event	37 (14·1%)	13 (4·9%)
Lack of efficacy	6 (2·3%)	32 (12·0%)
Administrative	9 (3·4%)	20 (7·5%)

*Data from table 3.

Table: **Reasons given in published report of American multicentre study[20] for withdrawal (censoring) for secondary end-point analysis***

to cyclosporin had been rescued with tacrolimus. The distorting roles of tacrolimus overdosage and a high rate of toxicity were clarified by separate analyses of the early (high dose) and late (reduced dose) phases of the trial. The statistical analysis, based on the intent-to-treat approach, showed significantly greater freedom from acute rejection, intractable acute rejection, and chronic rejection with tacrolimus.

The American report

Published report

The analysis of this trial[20] was claimed to be by intention to treat, implying that all 529 enrolled patients except the 64 who died contributed data during the full period of study for all stipulated end-points: patients' and graft survival, rejection, intractable rejection, need for retransplantation, steroid need, and muromonab-CD3 (OKT3) use. However, the only analyses actually done by intent-to-treat were patients' and graft survival, which were not significantly different in the two arms. Data on all other end-points were heavily censored. Censored data included those from a large portion of the 465 patients who survived throughout the year of the study bearing their original grafts or after successful retransplantation (table, which is derived from ref 20).

Thus, rejection that occurred after a patient, for example, experienced an adverse event leading to

withdrawal of treatment, or who was withdrawn from the trial for any reason, was not counted in the published analysis. Inexplicably, only 30 of the 64 deaths and 38 of the 52 retransplantations were in the censored list (table). The rescue role of tacrolimus in reducing the overall incidence of retransplantation was obliquely discussed: "The low number of second transplantations for refractory rejection may have been due, in part, to the effectiveness of tacrolimus in treating patients in the cyclosporine group who had refractory rejection". In fact, grafts rescued by tacrolimus accounted for 20 (9·5%) of the 210 surviving grafts credited by intent-to-treat analysis to the cyclosporin arm at the end of the year.

Inappropriate use of the Kaplan-Meier method compounded the problem of secondary end-point analysis. Instead of using all the data generated by the patients who lived through the duration of the study, end-points subsequent to censoring were obtained by life-table (Kaplan-Meier) calculations. An assumption underlying this calculation is that censoring is random with respect to treatment assignment.[21] Both the number of patients censored (102 cyclosporin *vs* 83 tacrolimus) and most of the reasons were differentially distributed in one treatment arm or the other (table). Censoring because of adverse events was more frequent in the tacrolimus arm, while that because of lack of efficacy and "administrative reasons" was more frequent in the cyclosporin arm.

Reanalysis

The impression was left by the published American account that the better efficacy and greater toxicity of tacrolimus essentially balanced each other,[22] which prompted us to reanalyse the original database, using intent-to-treat methodology throughout. A clearer picture emerged. As in the European trial, randomisation produced groups that were similar at the outset with respect to all important prognostic factors. With the "freedom from" formulation of the published study, several of its conclusions were confirmed. However, the numerical results and their statistical significance were different from published for all end-points except for the
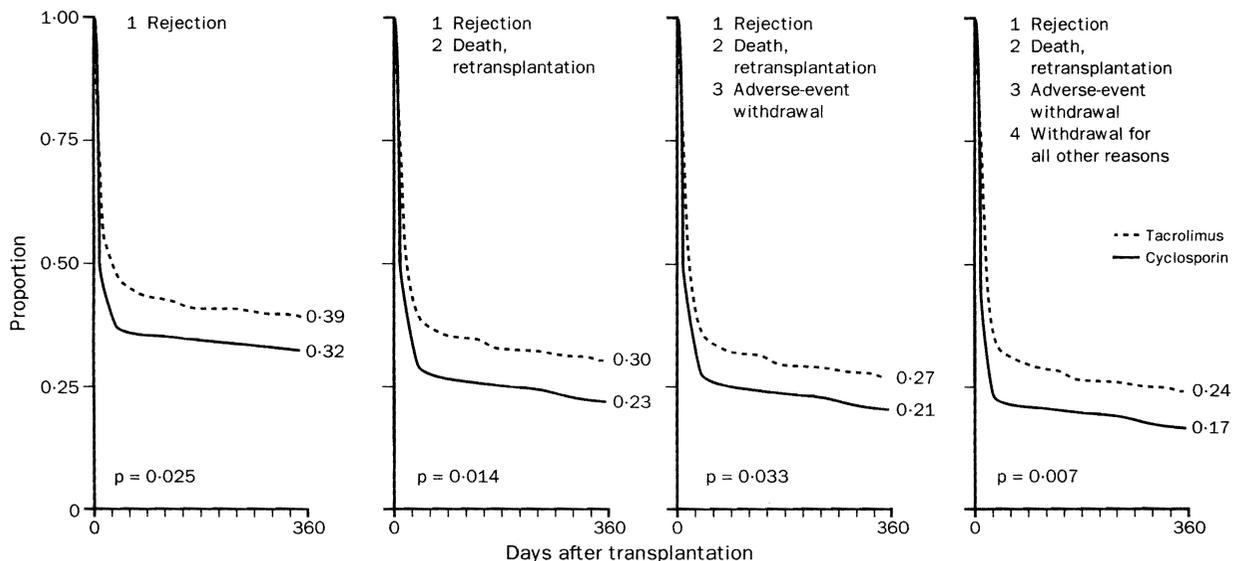

Figure 3: **Freedom from various undesirable end-points in reanalysis of American trial**
Log-rank test.

1-year patients' (88%) and graft (80·5%) survival. Consequently, reanalysis required stepwise restoration of all the censored categories shown in the table. On reanalysis, freedom from rejection as a single end-point was accomplished in 39% compared with 32% of the patients randomised to tacrolimus and cyclosporin, respectively (figure 3, left), which compares with 32% and 24% in the published report. When data were progressively reinserted from the different categories of censored patients in layers 2, 3, and 4 of the reanalysis, the "freedom from" curves of both arms progressively descended (figure 3). However, tacrolimus' superiority was maintained throughout.

In our reanalysis, restoration of freedom from adverse events had an almost immeasurable effect on the gap between tacrolimus and cyclosporin that was present before restoration. This corrected the impression left by the published report that tacrolimus' greater efficacy was balanced by increased toxicity. Thus the gap between tacrolimus and cyclosporin was similar for all the end-points in figure 3.

The most clinically relevant results of the reanalysis are shown in figure 4. After 1-year of follow-up, 98% of the patients randomised to tacrolimus were not diagnosed with refractory rejection compared with 87% in the competing arm. The log-rank p for all secondary end-points in the reanalysis indicated that experience of treatment failure was significantly more favourable with tacrolimus, as evaluated over the full year. The composite freedom at 1 year from the three factors that haunt transplant recipients (refractory rejection, retransplantation, and death) was 80% for tacrolimus and 70% for cyclosporin (figure 4, right).

Discussion

Regulatory agencies increasingly insist on controlled randomised trials as a prerequisite for marketing new drugs. This policy has powerful support in academic circles, for reasons that go beyond intellectual merit. Fiscal, administrative, and professional opportunities are generated within each component of the regulatory/
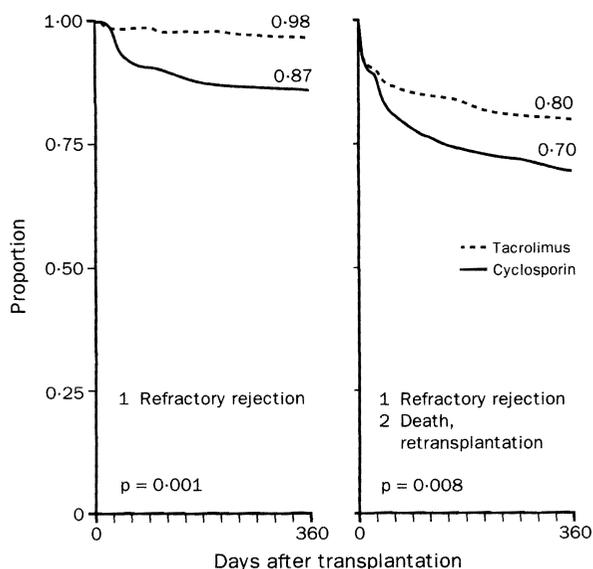
pharmaceutical/academic triad that drives such trials. The consequent range of possible conflicts of interest has made randomised trials a magnet for criticism.[17,23-25] The most damaging potential allegation has been that such studies are frequently carried out to obtain answers that are already known. We will not further labour the evidence that the tacrolimus/cyclosporin multicentre trials were a cruel and expensive example of this syndrome.

How were the "subjects" recruited for the investigation? Whether informed consent is possible in a captive population of patients, such as those on an organ-transplant waiting list, has been asked before.[17,26] Was anything worthwhile achieved by the multi-year detour off the developmental highway for tacrolimus that already had been cleared for the liver and all the other vital organ allografts by the spring of 1990? At journey's end, a talented group of clinician-investigators in twenty different centres had surmounted a learning curve for an experimental drug by simply reintroducing flexibility into a protocol that by the nature of transplantation biology can never be applied in the same way to any two recipients.

Did these randomised trials have an effect opposite to the objectives of improved and less expensive care of patients? Surprisingly, similar multicentre trials were mandated for other transplant indications, organ by organ. By the time these started, evidence had accrued from twenty-four US kidney-transplant centres with access to tacrolimus that the actuarial half-life of cadaver renal allografts was projected to be 14 years in recipients treated with maintenance tacrolimus versus 8 years with any previously available immunosuppressant, including cyclosporin.[27] If these projections prove valid, the cost savings lost to the taxpayer by the 5-year delay in use of tacrolimus for renal transplantation will have been immense.

Besides such ethical ramifications, we ask what influence the randomised trial mind-set is having on genuine clinical research, the atrophy of which has been mourned by Ahrens.[28] Tacrolimus was developed in Pittsburgh the old-fashioned way—funded by a National Institutes of Health grant (DK 29961) and a grant from the Veterans Administration. All clinical fees earned by faculty members of the University of Pittsburgh Transplantation Institute beyond their full-time salaries have always been invested in research rather than taken in personal income. This was a pure full-time system. Costs of hospital stay were defrayed in the usual ways.

No contractual financial agreement of any kind was ever signed, or even discussed, with representatives of Fujisawa, whose sole commitment was a gentleman's agreement to supply the investigative drug and up-to-date scientific information. The reward for the clean-hands policy was freedom from the kind of group-think decision-making that resulted in the obfuscating design used for the multicentre randomised trials. That experimental design ultimately was responsible for the American multicentre report that required reanalysis to be comprehensible.

Figure 4: **Freedom from refractory rejection and refractory rejection plus graft loss (from death or retransplantation) in reanalysis of American randomised trial**

References

1 Starzl TE, Todo S, Fung J, Demetris AJ, Venkataramanan R, Jain A. FK 506 for human liver, kidney, and pancreas transplantation. *Lancet* 1989; ii: 1000–04.

2 Fung JJ, Todo S, Jain A, et al. Conversion of liver allograft recipients with cyclosporine related complications from cyclosporine to FK 506. *Transplant Proc* 1990; **22:** 6–12.

3 Todo S, Fung JJ, Starzl TE, et al. Liver, kidney, and thoracic organ transplantation under FK 506. *Ann Surg* 1990; **212:** 295–305.

4 Starzl TE, Fung J, Jordan M, et al. Kidney transplantation under FK 506. *JAMA* 1990; **264:** 63–67.

5 Starzl TE, Iwatsuki S, Van Thiel DH, et al. Evolution of liver transplantation. *Hepatology* 1982; **2:** 614–36.

6 Todo S, Fung JJ, Starzl TE, et al. Single center experience with primary orthotopic liver transplantation under FK506 immunosuppression. *Ann Surg* 1994; **220:** 297–309.

7 Starzl TE, Marchioro TL, Waddell WR. The reversal of rejection in human renal homografts with subsequent development of homograft tolerance. *Surg Gynecol Obstet* 1963; **117:** 385–95.

8 Starzl TE, Putnam CW, Halgrimson CG, et al. Cyclophosphamide and whole organ transplantation in human beings. *Surg Gynecol Obstet* 1971; **133:** 981–91.

9 Starzl TE, Weil R III, Iwatsuki S, et al. The use of cyclosporin A and prednisone in cadaver kidney transplantation. *Surg Gynecol Obstet* 1980; **151:** 17–26.

10 Calne RY, Rolles K, White DJG, et al. Cyclosporin A initially as the only immunosuppressants in 34 recipients of cadaveric organs; 32 kidneys, 2 pancreases, and 2 livers. *Lancet* 1979; ii: 1033–36.

11 Fung JJ, Todo S, Tzakis A, et al. Conversion of liver allograft recipients from cyclosporine to FK506-based immunosuppression: benefits and pitfalls. *Transplant Proc* 1991; **23:** 14–21.

12 Starzl TE, Abu-Elmagd K, Tzakis A, Fung JJ, Porter K, Todo S. Selected topics on FK 506: with special references to rescue of extrahepatic whole organ grafts, transplantation of "forbidden organs", side effects, mechanisms, and practical pharmacokinetics. *Transplant Proc* 1991; **23:** 914–19.

13 Starzl TE, Hakala TR, Rosenthal JT, Iwatsuki S, Shaw BW Jr. Variable convalescence and therapy after cadaveric renal transplantation under cyclosporin A and steroids. *Surg Gynecol Obstet* 1982; **154:** 819–25.

14 Fung J, Abu-Elmagd K, Jain A, et al. A randomized trial of primary liver transplantation under immunosuppression with FK 506 vs cyclosporine. *Transplant Proc* 1991; **23:** 2977–83.

15 Jain A, Venkataramanan R, Cadoff E, et al. Effect of hepatic dysfunction and T-tube clamping on FK 506 pharmacokinetics and trough concentrations. *Transplant Proc* 1990; **22:** 57–79.

16 Abu-Elmagd K, Fung JJ, Alessiani M, et al. The effect of graft function on FK506 plasma levels, doses, and renal function, with particular reference to the liver. *Transplantation* 1991; **52:** 71–77.

17 Starzl TE. The puzzle people. Pittsburgh: University of Pittsburgh Press, 1992: 231–42.

18 Starzl TE. The puzzle people. Pittsburgh: Univeristy of Pittsburgh Press, 1992: 288–308.

19 The European FK506 Multicentre Liver Study Group. Randomised trial comparing tacrolimus (FK506) and cyclosporin in prevention of liver allograft rejection. *Lancet* 1994; **344:** 423–28.

20 The US Multicenter FK 506 Liver Study Group. A comparison of tacrolimus (FK 506) and cyclosporine for immunosuppression in liver transplantation. *N Engl J Med* 1994; **331:** 1110–15.

21 Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958; **53:** 457–81.

22 Calne RY. Immunosuppression in liver transplantation. *N Engl J Med* 1994; **331:** 1154–55.

23 The Standards of Reporting Trials Group: a proposal for structured reporting of randomized controlled trials. *JAMA* 1994; **272:** 1926–31.

24 Angell M. The Nazi hypothermia experiments and unethical research today. *N Engl J Med* 1990; **322:** 1462–64.

25 Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *N Engl J Med* 1994; **331:** 394–98.

26 Starzl TE. Protecting the patient's interest. *Kidney Int* 1985; **28:** S31–33.

27 Gjertson DW, Cecka M, Terasaki PI. The relative effects of FK 506 and cyclosporine on short- and long-term kidney graft survival. *Transplantation* (in press).

28 Ahrens EJ Jr. The crisis in clinical research: overcoming institutional obstacles. New York: Oxford University Press, 1992.